Arch Virol (2005) 150: 273–286
DOI 10.1007/s00705-004-0410-z

Archives of Virology
Printed in Austria

Nucleotide sequence, genome organization and phylogenetic analysis of Strawberry pallidosis associated virus, a new member of the genus *Crinivirus**

I. E. Tzanetakis[1], **J. Reed**[1], and **R. R. Martin**[2]

[1]Department of Botany and Plant Pathology and Center for Gene Research and Biotechnology, Oregon State University, Corvallis, Oregon, U.S.A.
[2]USDA-ARS, Horticultural Crops Research Laboratory, Corvallis, Oregon, U.S.A.

Received July 1, 2004; accepted August 16, 2004
Published online October 21, 2004

Summary. The complete nucleotide sequence of Strawberry pallidosis associated virus (SPaV), a newly identified member of the genus *Crinivirus*, family *Closteroviridae* has been determined. RNA 1 is 8067 nucleotides long and encodes at least three open reading frames (ORFs). The first ORF (ORF 1a) specifies a multifunctional protein that has papain-like proteinase, methyltransferase and RNA helicase domains. The RNA-dependent-RNA polymerase is encoded in ORF 1b and is probably expressed by a +1 ribosomal frameshift. The 3′ ORF of RNA 1 encodes a small protein with two potential transmembrane helices. RNA 2 is 7979 nucleotides long and encodes 8 ORFs, similar in amino acid sequence and arrangement with those of other criniviruses. SPaV encodes the largest structural protein of closteroviruses sequenced to date as the minor coat protein of the virus has molecular mass of approximately 80 kDa. The 3′ non-translated regions share nucleotide sequence identities of about 56% and the predicted folding of the non-translated regions is similar. Phylogenetic analyses reveal that SPaV is related most closely to *Abutilon yellows virus* and *Beet pseudo-yellows virus*, another virus that has been identified recently to cause identical symptoms on strawberry indicator plants as SPaV.

Introduction

The majority of commercially grown strawberry cultivars (*Fragaria ananassa*) exhibit no symptoms when infected with a single virus. Symptoms appear when

*The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession numbers AY488137 and AY488138 for RNA 1 and 2, respectively.

two or more of the about 15 viruses infecting strawberry [42] are found in a plant. The lack of symptoms on strawberry plants delayed virus identification and detection until recently [8]. Detection was carried out primarily by grafting suspicious material onto a variety of *Fragaria* indicator plants and observation of symptoms that developed on them [11]. It was not until recently that causal agents were identified and modern detection methods were developed for strawberry viruses [20, 26, 37, 43, 45, 47, 48, 54].

Strawberry pallidosis is a disease that was first identified in 1957 [50]. By definition, pallidosis disease (PD) is caused by graft transmittable agent(s) that cause marginal chlorosis and epinasty on *F. virginiana* indicator plants 'UC-10' and 'UC-11' while *F. vesca* plants remain asymptomatic. PD is asymptomatic in most commercial cultivars grown today but it is believed to act synergistically and exacerbate symptoms in plants infected with other viruses and especially *Strawberry mild yellow-edge virus* and *Strawberry mottle virus* [54]. PD recently has been associated with *Beet pseudo-yellows virus* (BPYV) [53] and Strawberry pallidosis associated virus (SPaV) [50], both members of the genus *Crinivirus*, family *Closteroviridae*, transmitted by the greenhouse whitefly, *Trialeurodes vaporariorum* [50].

Members of the family *Closteroviridae* have the largest genomes of all plant positive-strand RNA viruses approaching 20 Kb. The family has three genera, *Closterovirus, Ampelovirus* and *Crinivirus*. Species of the first two are monopartite while criniviruses have bipartite genomes. Closterovirus members are transmitted by aphids, ampelovirus by mealybugs and crinivirus by whiteflies.

This communication reports the complete nucleotide sequence and genome organization of SPaV. Also, an analysis of the conserved enzymatic domains as well as the structural proteins of SPaV and all other sequenced criniviruses is presented. Phylogenetic analysis of four open reading frames (ORFs) strongly suggests SPaV grouping within the genus *Crinivirus*.

Materials and methods

RNA purification, cDNA synthesis, cloning and SPaV genome amplification

A field isolate (M1) from Maryland, now deposited in the National Clonal Germplasm Repository (Corvallis, Oregon) as CFRA 9094 was utilized for determination of the complete genome of SPaV. Single-stranded (ss) and double-stranded (ds) RNA were extracted as described previously [18, 56].

For all reactions described, enzymes, plasmids and kits from Invitrogen Corp. (Carlsbad, CA) were utilized according to the manufacturer's protocols unless otherwise stated. For the complete genome, except the 5′ and 3′ termini, dsRNA was utilized as template for cDNA synthesis and reverse transcription-polymerase chain reaction (RT-PCR). Briefly, dried dsRNA extracted from the equivalent of 10 g of tissue was incubated with 20 mM methyl mercuric hydroxide in the presence of 0.5–1 μg of random hexameric nucleotide primers for 30 min at room temperature. Reverse transcription was performed utilizing the Thermoscript® reverse transcriptase. Second-strand synthesis was performed as described previously [20] and cDNA was cloned into a pCR 2.1 vector. Recombinant plasmids were assayed for inserts either by direct amplification of the fragments from bacterial colonies exploiting M13 forward and reverse primers or by plasmid purification and digestion with restriction endonucleases.

The sequence of SPaV was acquired utilizing reverse transcription-polymerase chain reaction (RT-PCR). Because of the lack of clones representing some regions of the genome the original cloning inserts were employed for development of oligonucleotide primers for PCR amplification. All primers were developed after the alignment of at least two clones. For amplification of all fragments, *Platinum*® Taq polymerase was used. The PCR program consisted of denaturation at 94 °C for 5 min, followed by 40 cycles of 45 sec at 94 °C, 30 sec at 55 °C and 2–4 min (depending on the putative fragment length) at 68 °C. A final extension step of 10 min at 68 °C terminated the program. The 3′ termini of both RNAs were determined by adding an oligo-adenosine tail [41] onto purified dsRNA and performing reverse transcription as described above with an oligo dT primer. For PCR amplification, a forward primer near the 3′ end of the known sequence, together with the oligo dT primer was used for PCR amplification. Each of the 3′ ends were amplified at least twice. For amplification of the 5′ termini the adenylated dsRNA was utilized and PCR amplification was achieved utilizing an oligo-thymidine primer as the reverse primer and two individual oligonucleotide primers derived from near the 3′ end of the sequence of the minus strands. Commercially available 5′ rapid amplification of cDNA ends (RACE) also was used according to the manufacturer's instructions using single-stranded RNA extracted according to the Hughes and Galau method [18]. Each of the 5′ amplifications were performed twice except for that of RNA 2 utilizing adenylated dsRNA as template that was amplified once.

Sequencing, genome and phylogenetic analysis

All fragments were amplified twice by RT-PCR and each individual reaction was cloned into the pCR 2.1 vector. The PCR products and two or more clones obtained from each reaction were sequenced in both orientations. The consensus sequence was acquired after alignment of the RT-PCR derived sequences and the original clones of the fragment where available employing the ClustalW software (European Bioinformatics Institute). Sequencing reactions were carried out at the facilities of the Central Services Laboratory at Oregon State University and Macrogen Inc. (Seoul, South Korea) using an ABI 377 and 3700 DNA sequencers, respectively.

The predictions of the secondary structure of the non-translated regions (NTR) of the virus were obtained with the mfold software [58] at default settings. Phylogenetic analysis was performed utilizing maximum parsimony on the PAUP* 4.0b 10 software [46]. Heuristic search applying ten replicates of random taxon sequence addition and the TBR (Tree Bisection Reconnection) swapping algorithm were the settings for the reconstruction of the phylograms. Bootstrap analysis consisted of 1000 replications utilizing the same parameters as above. Amino acid sequences of four genes of SPaV where chosen for the analysis: RNA-dependant-RNA polymerase (RdRp), heat shock protein 70 homolog (HSP70h), coat protein (CP) and minor coat protein (CPm). Members of all three genera of the family were utilized for the reconstruction of the HSP70h phylogram while for the other three phylograms, crinivirus genes were used. Analysis of the putative transmembrane domains of proteins and signal peptides were done using the respective tools at the CBS Prediction Servers. The identification of the putative ORFs was performed with both the ORF Finder and gene finding in viruses software. The conserved motifs of the proteins were identified applying the CDART (Conserved Domain Architecture Retrieval Tool) [13] and CDD (Conserved Domain Database) [27] programs.

Results and discussion

The complete nucleotide sequence of SPaV was determined. RNA 1 encodes the putative replication-related proteins and consists of 8067 nucleotides (nt). RNA 2 is 7979 nt-long and presumably encodes movement and structural proteins in

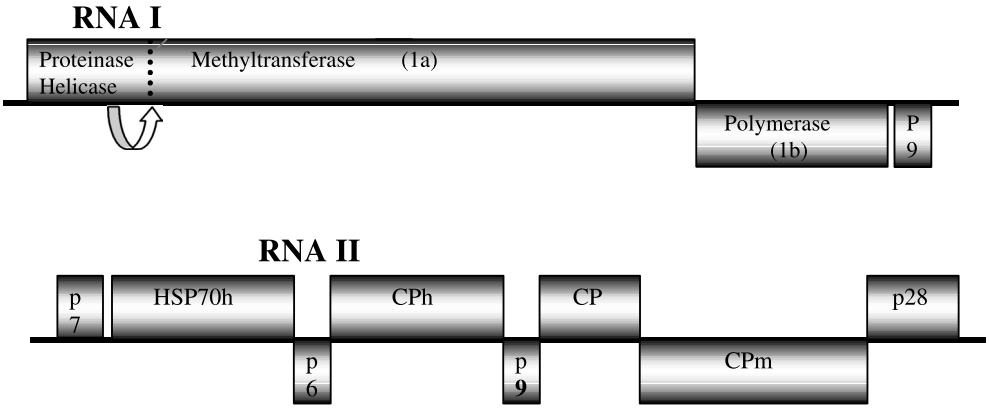

Fig. 1. Schematic representation of the genome organization of SPaV. The sizes of the ORFs are not to scale. *HSP70h* = heat shock 70 homolog, *CPh* = coat protein homolog, *CP* = coat protein, *CPm* = minor coat protein

addition to proteins of unknown function. The 5′ NTRs for RNA 1 and RNA 2 are 265 and 241 nt long, respectively. The nt sequence identity between these two regions is 46%, with no obvious similarities in the secondary structures of these RNAs (data not shown). SPaV is the only member of the genus that does not have the first six to ten nucleotides of the two RNAs identical [3, 16, 23, 25, 51]. Both 5′ termini have been determined by two tailing reactions with at least three PCR amplifications for each end and sequencing of more than eight clones in total. For RNA 2, the contig of the 5′ terminus also included two additional clones that were obtained from the original shotgun cloning. The 3′ NTR is 197 and 186 nucleotides long for RNA 1 and 2, respectively, and they share 56% nucleotide sequence identity. The predicted secondary structures of the two regions are similar although not identical (data not shown), and may serve as a signal for initiation for the viral polymerase [55].

The gene arrangement of SPaV is presented in Fig. 1. ORF 1a begins at nucleotide 266 and terminates at nucleotide 6106. The typical closterovirus polyprotein has one or two leader papain-like proteinases at the N′ terminus acting in cis and have been associated with accumulation of genomic RNA and long distance movement for *Beet yellows virus* (BYV), the type member of the *Closterovirus* genus [30, 33]. SPaV proteinase is also found at the N′ terminus of the polyprotein and is predicted to be the smallest such region of all criniviruses sequenced to date. The two catalytic residues [32] are a cysteine found at position 337 and the genus consensus consists of the motif αCWα where α stands for a hydrophobic residue (A, V, L, I, M, F, W or Y) while the other catalytic residue, histidine, is found at position 386. The putative cleavage site is found between the conserved glycine at position 404 and an isoleucine at position 405. The cleavage site is rather unusual for L-Pro orthologs; as they normally cleave between a Gly and Gly (or Ala or Ser). The leader proteinase of SPaV shows greater than 50% amino acid sequence similarity with the homologous domain of BPYV isolates. The putative molecular mass of this protein is estimated to be 47 kDa.

The methyltransferase (MTR) domain predicted to be between amino acids (aa) 457 and 815, is similar to that encoded by all members of the alphavirus superfamily [27]. The MTR motifs [38] are all well defined except for motif III. The crinivirus consensus is: XoXS/TaSS/NHXaXXAaRXC/TEN, motif I, ⟨12 aa⟩, aoaGGo, motif Ia1, ⟨9 aa⟩, ooaHaCXP, motif Ia2, XaDXKDXXR, motif II, ⟨20 aa⟩, CXXXXoXC, motif IIa1, ⟨9 aa⟩, VoVYDMT/SLXoMAXS/AaXXHK/GS/AK/RXaoαS/TXII, motif III, ⟨28 aa⟩, YXYGXXGEXaXHXXoXLooaL, motif IV, where α stands as above for a hydrophobic residue and o stands for a hydrophilic residue (E, D, Q, N, K, R or H) while X represents any amino acid.

A region of more than 800 amino acids separate the methyltransferase from the helicase domain of the protein. This region is the least conserved among the genus, except for two potential transmembrane helices that are found in all sequenced members of the genus. However, these potential transmembrane domains are not conserved at the family level (data not shown). In SPaV the two potential transmembrane helices are found between aa 1195–1217 and 1311–1333. The helicase domain is located at the 3′ terminus of ORF 1a between aa 1642 and 1946 and has six conserved motifs found in orthologous domains of the viruses that belong to the alphavirus superfamily. The first motif is the characteristic glycine-rich A site that binds the phosphate group followed by the B site that binds Mg^{2+} [7, 14, 27]. The genus consensus comprises of the following motifs: αXNKPPGAGKTTXIα, phosphate binding; ⟨58–59 aa⟩; XααCDEXFMαH, Mg^{2+} binding; ⟨16 aa⟩; XXαGDXNQIPα; ⟨28 aa⟩; XS/TYRCPXDX; ⟨77 aa⟩; XTVoEαQGG/CTFXXVXααR; ⟨13 aa⟩; QFαVXαSRH.

The RNA-dependent-RNA polymerase (RdRp) of closteroviruses presumably is expressed by a +1 ribosomal frameshift [2, 19, 22] which is unique among all positive-strand RNA viruses. The termination signal of 1a for SPaV is UUUGA (amber stop codon is underlined). All criniviruses sequenced to date have the two uridines before the stop codon which may indicate that they are involved in the +1 ribosomal frameshift although no particular structures favoring the slippage were identified utilizing the RNA folding software (data not shown) as was the case with all the other criniviruses [3, 23, 25]. The RdRp ORF (1b) is 505 codons encoding a protein of 58 kDa. The protein shows high amino acid sequence similarity with the sequenced RdRp genes of the other criniviruses reaching 68% aa identity and more than 80% aa similarity with that of the BPYV isolates. The eight conserved motifs of the positive-strand virus RdRps identified by Koonin [24] are present in SPaV and all crinivirus RdRps with consensus of: MaKGXXKPKaDX, motif I – possibly involved in RNA binding; ⟨10 aa⟩; XNIαYYooXαXαααSPαFLXXFXRα, motif II; ⟨7 aa⟩; αααYSGMNXXXLXXαα, motif III; ⟨14 aa⟩; EIDFXXFDKXQG, motif IV; ⟨42–43 aa⟩; GXQRRTGSPNTWLSNTLXTXXααLXXY, motif V involved in substrate binding; ⟨7 aa⟩; αLVSGDDSLI, motif VI, the active site of the enzyme (underlined GDD shows the triad Gly-Asp-Asp of the active site); ⟨26 aa⟩; SXPYFCSK, motif VII – possibly involved in RNA binding; ⟨10 aa⟩; VαPDXXRαFEK, motif VIII.

A small ORF that encodes a putative 79 amino acid, 9 kDa protein is found at the 3′ end of RNA 1 (Fig. 1). It starts at nt 7631 and terminates at nt 7870. This

putative protein has two potential transmembrane domains that span from amino acids 4–26 and 52–74. When the amino acid sequence is examined for putative signal peptides, it was found that there is a potential cleavage site between amino acids 21 and 22 and the program (SignalP V2.0.b2 at, http://www.cbs.dtu.dk/services/SignalP-2.0/) identified this as a signal peptide for bacteria and eukaryotes. This is the first protein with two putative transmembrane domains other than ORF 1a, identified to date in all the closteroviridae and the first potential signal peptide encoded by a positive strand RNA plant virus.

The 3′ proximal half of the closterovirus and ampelovirus genomes and RNA 2 for criniviruses encode genes that are expressed via subgenomic RNAs. Five of these genes (p7, HSP70h, CPh, CP and CPm) are involved in movement and assembly and are conserved in all members of the family. The first of these ORFs is a small peptide with a strong putative transmembrane helix motif that is involved in virus movement. In SPaV, the ORF encoding corresponding, ∼7 kDa putative peptide starts at the nt 242 and terminates at nt 430. The potential transmembrane helix spans aa 10–29. Recently, a cysteine at the N′ terminus of the homologous protein of BYV [35] has been shown to form disulfide bonds leading to the dimerization of the protein. The putative SPaV protein does not contain the Cys residue, which is not conserved across the members of the genus crinivirus. The role of the dimerization of the protein seems species specific since not all members of the closterovirus genus contain the residue at their N′ terminus.

The second ORF is the HSP70h gene of the virus. The protein has 556 residues and mass of 62 kDa. Heat shock proteins are molecular chaperones with an ATPase domain located at the N′ terminus of the protein and a substrate identification domain at the C′ terminus and accommodate correct folding of cellular proteins [6]. The closterovirus HSP70h are found in plasmodesmata, associated with virions and involved in virus movement [28, 29, 34, 39, 49]. All five motifs identified in the ATPase domain of HSP70h proteins [5] are present in the HSP70h of SPaV. The alignment of the conserved regions of the HSP70h genes of SPaV and other criniviruses gave the consensus: KαGLDFGTTFSTαS – motif I, phosphate binding domain ⟨148 aa⟩ FXXRRIαNEPSAAAαYXαSK/RX, motif II ⟨6 aa⟩ ααYDFGGGTFDXSLI motif III, phosphate binding domain ⟨13 aa⟩ GDSαLGGRDIDXXIXXXαXXXXXα – motif IV, adenosine binding domain ⟨70 aa⟩ and αααααVGGSSLLXXα motif V.

The next ORF codes for a small peptide of 6 kDa of unknown function. It begins after the termination signal of HSP70h at nt 2347 and terminates at nt 2508. The 53 amino acid peptide is similar with homologous proteins in BPYV and *Cucurbit yellow stunting disorder virus* (CYSDV) and the C′ terminus of *Little cherry virus-1* (LChV-1) HSP70h. The ORF has amino acid sequence identities that reach 40% with the homologous gene of BPYV, while the amino acid similarities exceed 50% among all three of these viruses for this ORF.

The next ORF is homologous with closterovirus proteins involved in virus movement and found as a structural protein in virions and it has been shown to have homologies with coat proteins at its' C′ terminus, making it a coat protein homolog (CPh) [30]. It was suggested that the conserved residues, arginine and aspartate

may be involved in stabilization of the structure of the proteins by forming salt bridges [10], however, Jagadish et al. (1993) using site directed mutagenesis showed this to be unlikely in potyviruses. The SPaV CPh is 518 aa long and has mass of 60 kDa. The putative start codon is at nt 2502, before the termination signal of p6 and it terminates at nt 4058. The protein is most closely related to the homologous peptide of BPYV with 47% and 67% amino acid sequence identity and similarity, respectively. In addition to the conserved arginine and aspartate identified by Napuli et al. [30] located at aa positions 434 and 471, respectively in SPaV, criniviruses encode for a conserved serine at position 340 (data not shown). The distance between the serine residue and the conserved arginine is longer than the distances between these residues in other coat proteins, making it uncertain whether this residue has the same function as in other coat protein genes. A small ORF of 9 kDa is located after the CPh. The peptide is found only in criniviruses among the members of the family. The putative 85 amino acid peptide shows no significant homology with other proteins in the database other than the homologous proteins of the other criniviruses.

The last two proteins that are involved in virus movement are the major and minor (CP and CPm, respectively) coat proteins of the viruses. CP protects 95% of the genome of the virus while CPm covers the remaining 5% at the 5′ end of the genomes of members of the family closteroviridae [36, 40, 57]. The major CP of the virus spans from nt 4287 to 5033 and encodes a protein of 248 amino acids with a molecular mass of 28 kDa. The three conserved residues of all filamentous virus CP genes [10] are found at positions 119 (serine), 166 (arginine) and 203 (aspartate). The protein also bears conserved motifs at the C′ terminus found in the coat proteins of members of the family closteroviridae [27]. The minor coat protein start codon precedes the stop codon of the CP gene and the ORF starts at nt 5002 and terminates at nt 7071, coding for the largest structural protein of all closteroviruses sequenced to date, since it encodes a protein with 689 aa with a mass of nearly 80 kDa. The conserved serine, arginine and aspartate are found at positions 563, 607 and 648, respectively, in the CPm.

The 3′-terminal ORF of SPaV RNA 2 could code for a 28 KDa protein that starts at nt 7074 and terminates at nt 7793 encoding for 239 amino acids. This protein is one of the most diverse of the crinivirus encoded proteins, and the SPaV protein shows the greatest similarity with that of the BPYV with 27% amino acid sequence identity and more than 57% similarity.

The phylogenetic analysis of the HSP70h gene clearly places SPaV in the genus *Crinivirus* (Fig. 2A). The CP phylogram (Fig. 2B) reveals a closer relationship of SPaV with *Abutilon yellows virus* (AbYV) (CP is the only completely sequenced gene of AbYV) than BPYV, while phylogenetic analysis of the HSP70h, polymerase, and CPm (data not shown for the latter two) show the clustering of SPaV with BPYV. These three viruses form a distinct cluster apart from all other members of the genus.

The complete nucleotide sequence of SPaV, a newly identified member of the family *Closteroviridae* has been determined. The virus encodes genes found in all members of the family and phylogenetic analysis demonstrated that the virus is

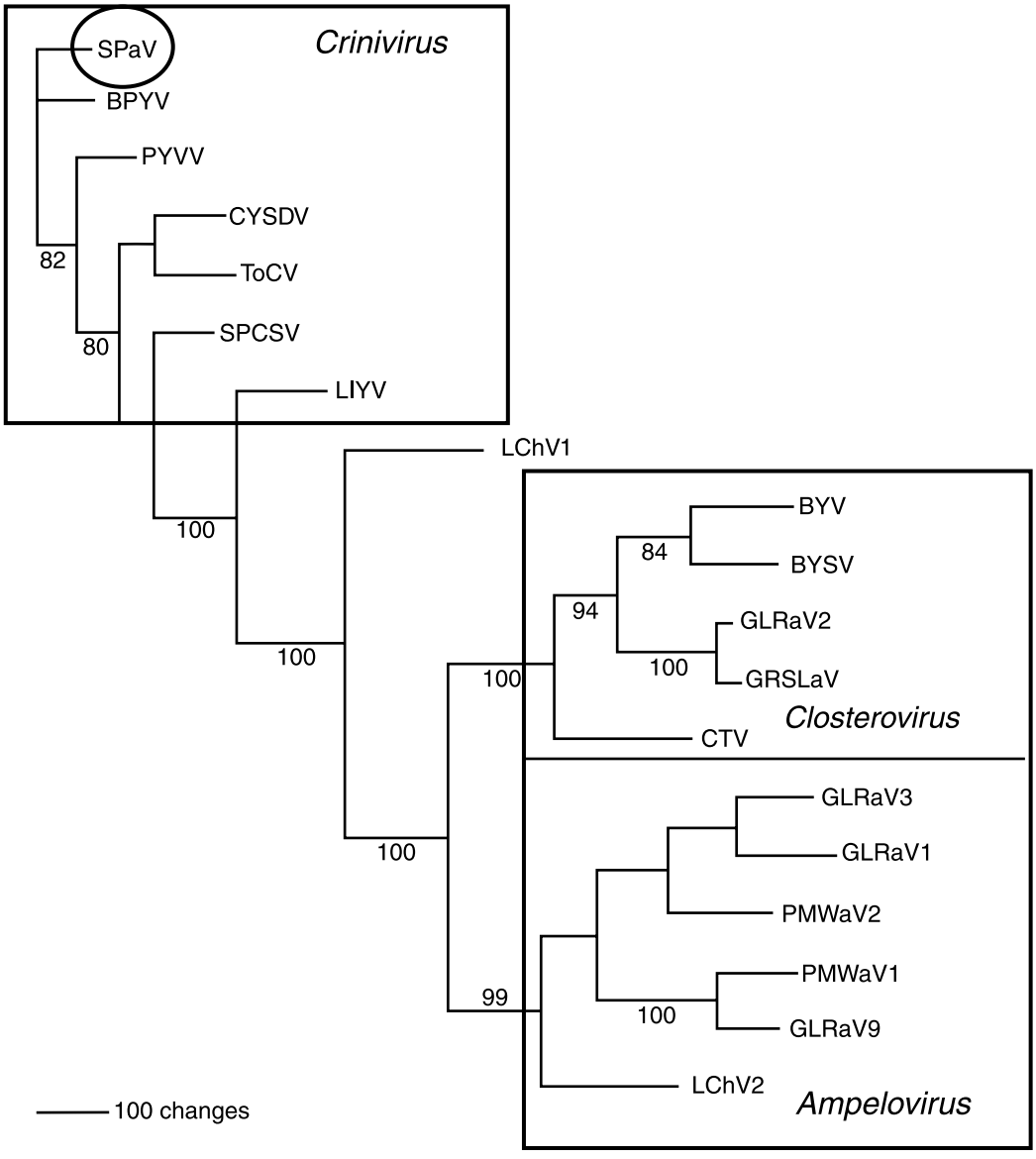

Fig. 2A. Phylogram of heat shock protein 70 homolog of SPaV and other closteroviruses. Abbreviations and GenBank accession numbers: BPYV, *Beet pseudo-yellows virus*, AAQ97386; BYSV, *Beet yellow stunt virus*, AAC55662; BYV, *Beet yellows virus*, NP041872; CTV, *Citrus tristeza virus*, NP042864; CYSDV, *Cucurbit yellow stunting disorder virus*, NP851572; GLRV1, *Grapevine leafroll associated virus-1*, AAK38612; GLRV2, *Grapevine leafroll associated virus-2*, AAR21242; GLRV3, *Grapevine leafroll associated virus-3*, NP813799; GRSLV, *Grapevine rootstock stem lesion associated virus* NP835247; LChV-1, *Little cherry virus-1*, NP045004; LChV2, *Little cherry virus-2*, AF531505; LIYV, *Lettuce infectious yellows virus*, NP619695; PMWaV1, *Pineapple mealybug wilt-associated virus-1*, AAL66711; PMWaV2, *Pineapple mealybug wilt-associated virus-2*, AAG13941. SPaV, *Strawberry pallidosis associated virus*, AAO92347; SPCSV, *Sweet potato chlorotic stunt virus*, NP689401; ToCV, *Tomato chlorosis virus*, AF024630. Bootstrap values are shown as percentage value and only the nodes over 50% are labeled. The bar represents 100 amino acid changes over the length of the proteins

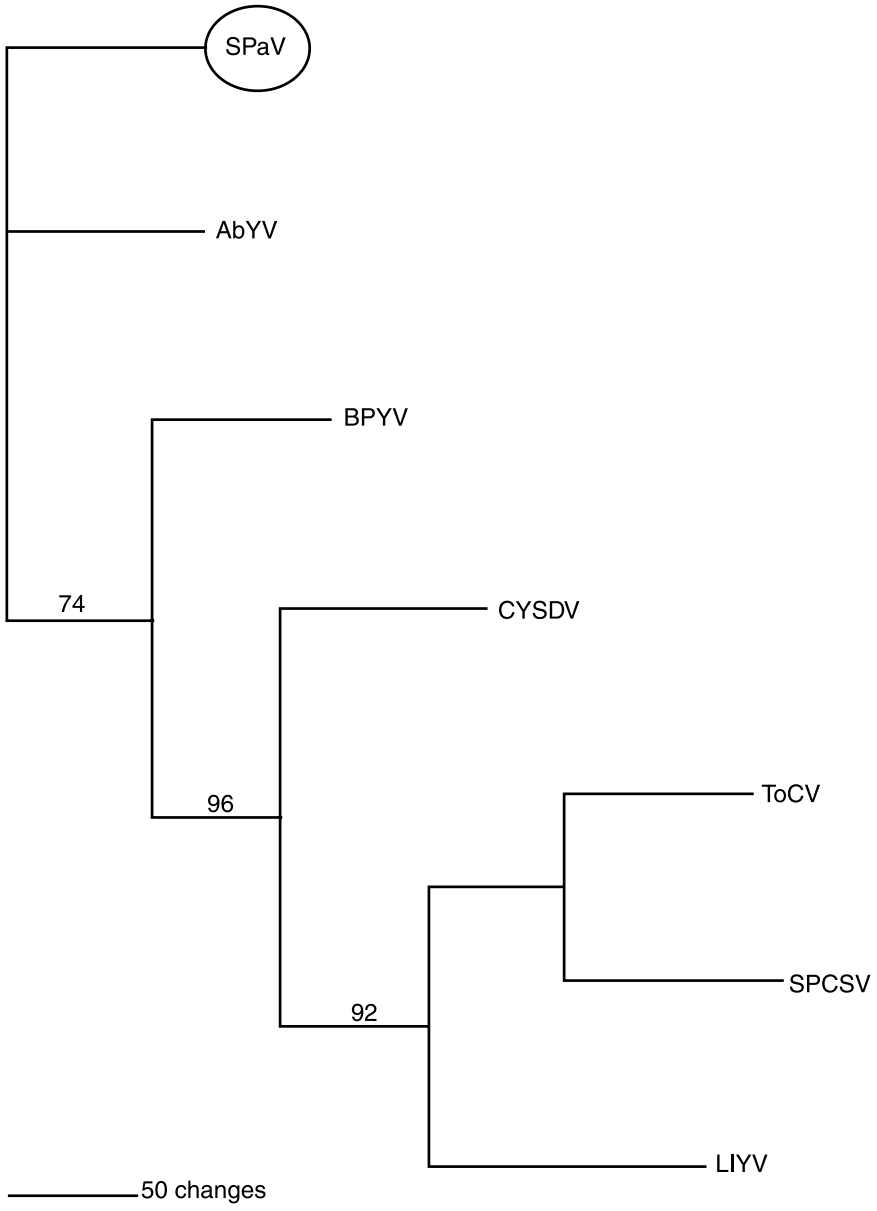

Fig. 2B. Phylogram of the coat protein of SPaV and other criniviruses. Abbreviations and GenBank accession numbers: AbYV, *Abutilon yellows virus*, AAR00224; BPYV, *Beet pseudo-yellows virus*, NP940792; CYSDV, *Cucurbit yellow stunting disorder virus*, NP851576; LIYV, *Lettuce infectious yellows virus*, NP619697; SPaV, *Strawberry pallidosis associated virus*, AAO92342; SPCSV, *Sweet potato chlorotic stunt virus*, NP689404; ToCV, *Tomato chlorosis virus*, AAR15080. Bootstrap values are shown as percentage value and only the nodes over 50% are labeled. The bar represents 50 amino acid changes over the length of the proteins

a member of the criniviruses. SPaV has some unique features not found in other members of the genus including; sequence differences at the very 5′ termini of RNA 1 and RNA 2, ORF 2 on RNA 1 has two potential transmembrane domains

and a potential signal peptide. Although unusual among viruses with segmented genomes there are examples where the nucleotides of the 5′ termini do not match [12, 15]. The two putative transmembrane domains found in the middle of ORF 1a is another feature that has been identified in SPaV and other members of the genus, while it was absent in the closteroviruses, BYV and *Citrus tristeza virus* (CTV). However, CTV possesses two similar putative transmembrane domains in the first of the leader proteinases. The role of the potential transmembrane domains in SPaV is unknown but it can be speculated that they may anchor the replication complex in cell membranes [9]. The putative ribosomal frame shift by which the virus polymerase is expressed has the very similar sequence before the putative +1 frameshift that has been identified in all criniviruses sequenced to date. As in the case of other criniviruses, computer assisted analysis in the region surrounding the site did not reveal any distinct structures that may facilitate this function. Additionally, no subgenomic RNA corresponding to the size expected if the RdRp was expressed from a subgenomic RNA was detected for either CYSDV or *Sweet potato chlorotic stunt virus* [3, 25]. Also, a start codon is found 12 nucleotides downstream from the putative frameshift site but no internal ribosome entry site that could initiate translation of the RdRp was identified (data not shown). SPaV is the only crinivirus that has been identified to encode a peptide with two putative transmembrane domains. This peptide also bears features similar to those of signal peptides with a putative cleavage site inside the domain that is found inside the membrane, making the possibility of actual cleavage less likely. RNA 2 of the criniviruses is more conserved than RNA 1. RNA 2 of SPaV contains the seven genes found in all other criniviruses that have been fully sequenced; p7, HSP70h, CPh, p9, CP, CPm and p28. In addition, SPaV contains a small ORF that follows the HSP70h which is also found in BPYV and CYSDV. The similarity of the peptide with the C′ terminus to the LChV-1 HSP70h suggests that the protein may have been part of a larger HSP70h gene that has evolved to an individual gene. The C′ terminus of heat shock proteins are involved in substrate recognition although in the case of *Beet yellows virus* such a function was not identified [1]. The other genes of RNA 2 (p6, p9 and p28) resemble homologous genes of other criniviruses in size and location.

Multiple attempts were performed for northern analysis and study of the expression pattern of the SPaV ORFs. SPaV has very low titer in strawberry [50]. More than ten RNA radioactive probes were used in Northern analysis and only the RNA corresponding to the CP ORF gave a faint signal (data not shown). In order to verify that the methodology was performed correctly experiments with BYV were carried out in parallel [36]. The BYV Northerns worked well even at a 20-fold dilution of RNA concentration to that used for SPaV detection, suggesting that the lack of signal in the SPaV assay was due to the low titer of the virus rather than a problem with the technique.

The phylogram of the coat protein (Fig. 2B) indicated that SPaV, BPYV and AbYV may belong to the same subgroup within the genus since they form a tight cluster separate from the other four species, although more sequences from more genes are needed to confirm this observation.

Criniviruses have become an emerging problem in world agriculture due to the movement of plant materials and naturalization or increasing range of the whitefly vectors. Small fruits are a new group of crops that have been identified to be infected with this group of viruses [52–54]. Recently, the greenhouse whitefly has become naturalized in many areas in the southern U.S. and become a pest in southern strawberry production areas of the U.S. [4]. The ability to transmit both viruses that are associated with the pallidosis disease may be the reason this virus disease has become widespread in these areas [17, 54].

Unlike other plant virus families which have a set number and arrangement of proteins encoded in the genome, family *Closteroviridae* shows great diversity in the genome size and genes encoded by its members ranging from that of BPVV [16, 51] and SPaV to that of *Citrus tristeza virus* [22]. This diversity within the family makes it essential to study the genomes of closteroviruses in order to better understand the evolution of the family, including the genome segmentation and the different transmission modes of the three genera.

Acknowledgments

The authors would like to thank Dr. Valerian Dolja for reviewing the manuscript and helpful suggestions. This project was funded by the North American Strawberry Growers' Association and the United States Department of Agriculture.

References

1. Agranovsky AA, Folimonova SY, Folimonov AS, Denisenko ON, Zinovkin RA (1997) The beet yellows closterovirus p65 homologue of HSP70 chaperones has ATPase activity associated with its conserved N-terminal domain but does not interact with unfolded protein chains. J Gen Virol 78: 535–542
2. Agranovsky AA, Koonin EV, Boyko VP, Maiss E, Frotschl R, Lunina NA, Atabekov JG (1994) Beet yellows closterovirus: complete genome structure and identification of a leader papain-like thiol protease. Virology 198: 311–324
3. Aguilar JM, Franco M, Marco CF, Berdiales B, Rodriguez-Cerezo E, Truniger V, Aranda MA (2003) Further variability within the genus *Crinivirus*, as revealed by determination of the complete RNA genome sequence of Cucurbit yellow stunting disorder virus. J Gen Virol 84: 2555–2564
4. Bi JL, Toscano NC, Ballmer GR (2002) Seasonal population dynamics of the greenhouse whitefly Trialeurodes vaporariorum (Homoptera: Aleyrodidae) on strawberries in Southern California. J Econ Entomol 95: 1179–1184
5. Bork P, Sander C, Valencia A (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. Proc Natl Acad Sci USA 89: 7290–7294
6. Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 chaperone machines. Cell 92: 351–366
7. Candresse T, Morch MD, Dunez J (1990) Multiple alignment and hierarchical clustering of conserved amino acid sequences in the replication associated proteins of plant RNA viruses. Res Virol 141: 315–329
8. Converse RH (1992) Modern approaches to strawberry virus research. Acta Horticult 308: 19–30

9. den Boon JA, Chen J, Ahlquist P (2001) Identification of sequences in Brome mosaic virus replicase protein 1a that mediate association with endoplasmic reticulum membranes. J Virol 75: 12370–12381
10. Dolja VV, Boyko VP, Agranovsky AA, Koonin EV (1991) Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation. Virology 184: 79–86
11. Frazier NW (1974) Detection of graft-transmissible diseases in strawberry by a modified leaf grafting technique. Plant Dis Rept 58: 203–207
12. Ge X, Scott SW, Zimmerman MT (1997) The complete sequence of the genomic RNAs of spinach latent virus. Arch Virol 142: 1213–1226
13. Geer LY, Domrachev M, Lipman DJ, Bryant SH (2002) CDART: protein homology by domain architecture. Genome Res 12: 1619–1623
14. Gorbalenya AE, Koonin EV (1989) Viral proteins containing the purine NTP-binding sequence pattern. Nucleic Acids Res 17: 8413–8440
15. Han SS, Karasev AV, Ieki H, Iwanami T (2002) Nucleotide sequence and taxonomy of Cycas necrotic stunt virus. Arch Virol 147: 2207–2214
16. Hartono S, Tomohide N, Genda Y, Okuda S (2003) Nucleotide sequence and genome organization of Cucumber yellows virus, a member of the genus Crinivirus. J Gen Virol 84: 1007–1012
17. Hokanson SC, Martin RR, Heflebower RF Jr, Rouse R, Maas J (2000) Survey of strawberry viruses occurring in commercial plantings in the state of Maryland, USA. Advances in Strawberry Research 18: 25–32
18. Hughes DW, Galau G (1988) Preparation of RNA from cotton leaves and pollen. Plant Molec Biol Reptr 6: 253–257
19. Jelkmann W, Fechtner B, Agranovsky AA (1997) Complete genome structure and phylogenetic analysis of little cherry virus, a mealybug-transmissible closterovirus. J Gen Virol 78: 2067–2071
20. Jelkmann W, Maiss E, Martin RR (1992) The nucleotide sequence and genome organization of strawberry mild yellow edge-associated potexvirus. J Gen Virol 73: 475–479
21. Jelkmann W, Martin RR, Maiss E (1989) Cloning of four plant viruses from small quantities of dsRNA. Phytopathology 79: 1250–1253
22. Karasev AV, Boyko VP, Gowda S, Nikolaeva OV, Hilf ME, Koonin EV, Niblett CL, Cline K, Gumpf DJ, Lee RF et al. (1995) Complete sequence of the citrus tristeza virus RNA genome. Virology 208: 511–520
23. Klaassen VA, Boeshore ML, Koonin EV, Tian T, Falk BW (1995) Genome structure and phylogenetic analysis of Lettuce infectious yellows virus, a whitefly-transmitted, bipartite closterovirus. Virology 208: 99–110
24. Koonin EV (1991) The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. J Gen Virol 72: 2197–2206
25. Kreuze JF, Savenkov EI, Valkonen JP (2002) Complete genome sequence and analyses of the subgenomic RNAs of *Sweet potato chlorosis stunt virus* reveal several new features for the genus *Crinivirus*. J Virol 76: 9260–9270
26. Lamprecht S, Jelkmann W (1997) Infectious cDNA clone used to identify strawberry mild yellow edge-associated potexvirus as causal agent of the disease. J Gen Virol 78: 2347–2353
27. Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH (2003)

CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Res 31: 383–387
28. Medina V, Peremyslov VV, Hagiwara Y, Dolja VV (1999) Subcellular localization of the HSP70-homolog encoded by beet yellows closterovirus. Virology 260: 173–181
29. Napuli AJ, Falk BW, Dolja VV (2000) Interaction between HSP70 homolog and filamentous virions of the Beet yellows virus. Virology 274: 232–239
30. Napuli AJ, Alzhanova DV, Doneanu CE, Barofsky DF, Koonin EV, Dolja VV (2003) The 64-kilodalton capsid protein homolog of Beet yellows virus is required for assembly of virion tails. J Virol 77: 2377–2384
31. Peng CW, Dolja VV (2000) Leader proteinase of the beet yellows closterovirus: mutation analysis of the function in genome amplification. J Virol 74: 9766–9770
32. Peng CW, Peremyslov VV, Mushegian AR, Dawson WO, Dolja VV (2001) Functional specialization and evolution of the leader protease of the family Closteroviridae. J Virol 75: 12153–12160
33. Peng CW, Napuli AJ, Dolja VV (2003) Leader proteinase of beet yellows virus functions in long-distance transport. J Virol 77: 2843–2849
34. Peremyslov VV, Hagiwara Y, Dolja VV (1999) HSP70 homolog functions in cell-to-cell movement of a plant virus. Proc Natl Acad Sci USA 96: 14771–14776
35. Peremyslov VV, Pan YW, Dolja VV (2004a) Movement protein of a closterovirus is a type III integral transmembrane protein localized to the endoplasmic reticulum. J Virol 78: 3704–3709
36. Peremyslov VV, Andreev IA, Prokhnevsky AI, Duncan GH, Taliansky ME, Dolja VV (2004b) Complex molecular architecture of beet yellows virus particles. Proc Natl Acad Sci USA 101: 5030–5035
37. Petrzik K, Benes V, Mraz I, Honetslegrova-Franova J, Ansorge W, Spak J (1998) Strawberry vein banding virus-definitive member of the genus Caulimovirus. Virus Genes 16: 303–305
38. Rozanov MN, Koonin EV, Gorbalenya AE (1992) Conservation of the putative methyltransferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand RNA viruses. J Gen Virol 73: 2129–2134
39. Satyanarayana T, Gowda S, Mawassi M, Albiach-Marti MR, Ayllon MA, Robertson C, Garnsey SM, Dawson WO (2000) Closterovirus encoded HSP70 homolog and p61 in addition to both coat proteins function in efficient virion assembly. Virology 278: 253–265
40. Satyanarayana T, Gowda S, Ayllon MA, Dawson WO (2004) Closterovirus bipolar virion: evidence for initiation of assembly by minor coat protein and its restriction to the genomic RNA 5′ region. Proc Natl Acad Sci USA 101: 799–804
41. Sippel AE (1973) Purification and characterization of adenosine triphosphate: ribonucleic adenyltransferase from *Escherichia coli*. Eur J Biochem 37: 31–40
42. Spiegel S, Martin RR (1998) Virus and viruslike diseases. In: Maas JL (ed) Compedium of strawberry diseases. American Phytopathological Society Press, St Paul, Minessota, pp 62–72
43. Spiegel S, Martin RR, Leggett F, ter Borg M, Postman J (1993) Characterization and geographical distribution of a new virus from Fragaria chiloensis. Phytopathology 83: 991–995
44. Stenger DC, Mullin RH, Morris TJ (1987) Characterization and detection of the strawberry necrotic shock isolate of tobacco streak virus. Phytopathology 77: 1330–1337
45. Stenger DC, Mullin RH, Morris TJ (1988) Isolation, molecular cloning, and detection of strawberry vein banding virus DNA. Phytopathology 78: 154–159
46. Swofford DL (2001) PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods) Version 4, Sinauer Associates, Sunderland, Massachusetts

47. Thompson JR, Wetzel S, Klerks MM, Vaskova D, Schoen CD, Spak J, Jelkmann W (2003) Multiplex RT-PCR detection of four aphid-borne strawberry viruses in *Fragaria* spp in combination with a plant mRNA specific internal control. J Virol Methods 111: 85–93
48. Thompson JR, Leone G, Lindner JL, Jelkmann W, Schoen CD (2002) Characterization and complete nucleotide sequence of *Strawberry mottle virus*: a tentative member of a new family of bipartite plant picorna-like viruses. J Gen Virol 83: 229–239
49. Tian T, Rubio L, Yeh H-H, Crawford B, Falk BW (1999) Lettuce infectious yellows virus: in vitro acquisition analysis using partially purified virions and the whitefly Bemisia tabaci. J Gen Virol 80: 1111–1117
50. Tzanetakis IE, Halgren AB, Keller KE, Hokanson SC, Maas JL, McCarthy PL, Martin RR (2004) Identification and detection of a virus associated with strawberry pallidosis disease. Plant Dis 88: 383–390
51. Tzanetakis IE, Martin RR (2004) Complete nucleotide sequence of a strawberry isolate of Beet pseudoyellows virus. Virus Genes 28: 239–246
52. Tzanetakis IE, Gergerich R, Fernandez G, Pesic Z, Martin RR (2003a) A crinivirus associated with a chlorotic line pattern in blackberry. Phytopathology 93: S85
53. Tzanetakis IE, Wintermantel WM, Martin RR (2003b) First report of Beet pseudo yellows virus in strawberry in the USA: A second crinivirus able to cause pallidosis disease. Plant Dis 87: 1398
54. Tzanetakis IE (2004) Molecular characterization of criniviruses and ilarviruses infecting strawberry. PhD dissertation submitted to Oregon State University, USA
55. Yi M, Lemon SM (2003) Structure-function analysis of the 3′ stem-loop of hepatitis C virus genomic RNA and its role in viral RNA replication. RNA 9: 331–345
56. Yoshikawa N, Converse RH (1990) Strawberry pallidosis disease: distinctive dsRNA species associated with latent infections in indicators and in diseased strawberry cultivars. Phytopathology 80: 543–548
57. Zinovkin RA, Jelkmann W, Agranovsky AA (1999) The minor coat protein of beet yellows closterovirus encapsidates the 5′ terminus of RNA in virions. J Gen Virol 80: 269–272
58. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415

Author's address: Dr. Robert R. Martin, USDA-ARS, Horticultural Crops Research Laboratory, 3420 NW Orchard Str., OR 97330 Corvallis, U.S.A.; e-mail: Martinrr@science.oregonstate.edu